MODEL BUILDING IN THE STOCK MARKET: SOME EXAMPLES FROM THE LITERATURE

William Warren Wallace

[illegible]

Library
Naval Postgraduate School
Monterey, California 93940

INTERNALLY DISTRIBUTED

REPORT

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

MODEL BUILDING IN THE STOCK MARKET:
SOME EXAMPLES FROM THE LITERATURE

by

William Warren Wallace

Thesis Advisor: Edward A. Brill

September 1973

T157008

Approved for public release; distribution unlimited.

INTERNALLY DISTRIBUTED

REPORT

Model Building in the Stock Market:

Some Examples from the Literature

by

William Warren Wallace
Lieutenant Commander, Supply Corps, United States Navy
B.S., New Mexico State University, 1962

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1973

Library
Naval Postgraduate School
Monterey, California 93940

ABSTRACT

The operations research technique of mathematical model building and testing as it has been applied in the field of stock market research is examined using various examples from the literature. The vehicle used for this examination is a general model for stock valuation. This general model illustrates that the value of a share of stock depends on estimates of future values of earnings or dividends, rates of return, and future stock prices. Therefore models are illustrated which relate to each of these variables.

This examination of past applications of model building in the stock market is intended to stimulate interest in this area of research, and toward this end some areas for future research are suggested.

TABLE OF CONTENTS

I. INTRODUCTION ---------- 4

II. A GENERAL STOCK VALUATION MODEL ---------- 5

III. EARNINGS ESTIMATION MODELS ---------- 16

IV. RATE OF RETURN ESTIMATION MODELS ---------- 22

V. PRICE ESTIMATION MODELS ---------- 32

A. THE RANDOM WALK HYPOTHESIS ---------- 32

B. PRICE MODELS ---------- 35

VI. SUMMARY ---------- 61

LIST OF REFERENCES ---------- 64

INITIAL DISTRIBUTION LIST ---------- 66

FORM DD 1473 ---------- 67

I. INTRODUCTION

This thesis is an effort to give a reader with a background in operations research an insight into various modelling techniques which have been used in the stock market, using a stock valuation model as a vehicle. With this concept in mind, the thesis does not attempt a comprehensive survey of stock valuation models, nor does it attempt to criticize the models which are presented.

The basic intention is to present examples of modelling techniques that are representative of the work which has been done in the past in order to point out, and hopefully stimulate interest in, a field of application for operations research techniques that has not received the publicity that has been accorded to other applications of operations research.

The method of exposition is to first present a general purpose stock valuation model and examine its basic components. Using this background each component is then examined in turn and models related to each component are presented. In this manner a wide range of models may be discussed and yet all are presented in the context of their relation to the general model.

II. A GENERAL STOCK VALUATION MODEL

The basic idea of stock valuation is to arrive at some sort of discounted present value for a share of stock. Although it is obvious that an investor's concepts of risk and utility should have some bearing on the value of a share of stock, the models discussed in this thesis do not explicitly include these concepts. The literature concerning risk and utility in the stock market is too vast to be included in the scope of this thesis. For some insight into the treatment of risk in stock valuation, Robichek (Ref. 1) has written a comprehensive article on the subject and has included an extensive bibliography. Also, Myers (Ref. 2) has proposed a very interesting treatment of risk in the form of a time-state-preference model.

The various stock valuation models which have been formulated generally differ in what is to be discounted and what horizon is to be used. Sloane and Reisman (Ref. 3) have proposed a general model which integrates several previous hypotheses regarding discounting and methods for financing growth. This model has been selected as the basic model for the discussion of stock valuation due to the good overview it gives of the discounting approach, and it will form the background against which the other models will be presented.

Sloane and Reisman classify the various approaches to discounting into the following five categories:

1. Discounting of anticipated earnings to perpetuity.
2. Discounting of anticipated earnings to a terminal date and discounting of anticipated terminal price.
3. Discounting of anticipated dividends to be received to perpetuity.
4. Discounting of anticipated dividends to a terminal date and discounting of anticipated terminal price.
5. Discounting of anticipated net cash flow.

They also specify four methods of financing growth by a firm:

1. Outside equity.
2. Retained earnings.
3. Purchase of new equity shares by existing stockholders.
4. Debt.

The following terminology is used in developing the general model:

V_o = value of investment at time zero

E_j = earnings in period j

k = yield or rate of discount of initial investors

j = general time period subscript

P_n = value of investment liquidated in period n

D_j = dividend in period j

I_o = equity available to firm at time zero

I_j^R = investment through retained earnings

r_j = rate of return on assets by firm in period j

k_j = earnings yield acquired on stock purchased by investors subsequent to time zero

I_j^S = value of new stock sold to new and prior stockholders in period j

$I_j^{S^N}$ = portion of I_j^S purchased by new stockholders

$I_j^{S^O}$ = portion of I_j^S purchased by old stockholders

I_j^D = amount of debt issued in period j

i_j = weighted average of interest rate paid on debt in period j

i = general time period subscript

It should be noted that the model uses a discount rate of investors, k, as well as the internal rate of return, r, and the interest paid on debt, i. These rates are not independently set and their functional relationships are not resolved by the model. The model will determine the value of any one of the variables once the other two have been assigned specified values, but it does not specify the value of any one independently of the others.

The general model is constructed by beginning with the first of the previously listed classifications of discounting hypotheses and adding the others on one at a time. The "discounting of earnings only" approach hypothesizes that the value of a stock is the present value of all earnings anticipated to perpetuity, discounted at an appropriate rate. That is,

$$V_o = \frac{E_1}{(1+k)^1} + \frac{E_2}{(1+k)^2} + \ldots + \frac{E_\infty}{(1+k)^\infty}$$

or

$$V_o = \sum_{j=1}^{\infty} E_j(1+K)^{-j}$$

The "earnings with finite horizon approach" is used when the investor intends to sell the stock at some time in the future. This approach hypothesizes that the value of a stock is the present value of all earnings from time zero until the stock is sold plus the present value of the price received at the time of the sale. That is,

$$V_o = \frac{E_1}{(1+K)^1} + \frac{E_2}{(1+K)^2} + \ldots + \frac{E_n}{(1+K)^n} + \frac{P_n}{(1+K)^n}$$

or

$$V_o = \sum_{j=1}^{n} \left[E_j(1+K)^{-j}\right] + P_n(1+K)^{-n}$$

The "dividends only" approach reasons that investors do not actually receive earnings. They receive dividends instead and therefore the dividends should be discounted, not the earnings. If the dividends are assumed to be received at the end of each period, then the value of a stock is the present value of all dividends discounted to perpetuity. That is,

$$V_o = \frac{D_1}{(1+k)^1} \quad \frac{D_2}{(1+k)^2} + \ldots + \frac{D_\infty}{(1+K)^\infty}$$

or

$$V_o = \sum_{j=1}^{\infty} D_j (1+k)^{-j}$$

However, in the absence of new outside equity or new debt financing, the dividends in any period j are the earnings for that period less the retained earnings. That is,

$$D_j = E_j - I_j^R \quad .$$

Therefore, the value of the stock may be expressed in terms of earnings and retained earnings as follows:

$$V_o = \frac{E_1 - I_1^R}{(1+K)^1} + \frac{E_2 - I_2^R}{(1+K)^2} + \ldots + \frac{E_\infty - I_\infty^R}{(1+K)^\infty}$$

By combining terms and rearranging,

$$V_o = \sum_{j=1}^{\infty} \left[E_j (1+K)^{-j} \right] - \sum_{j=1}^{\infty} \left[I_j^R (1+k)^{-j} \right]$$

This equation allows for the retention of earnings in any given period. Also, this equation can be made more general by explicitly taking into account the rate of return r, and the assets available to a firm in any period I_j. The earnings of a firm in any period are equal to the assets available in that period times the rate of return in that period. That is,

$$E_j = I_j r_j \quad .$$

Since I_j is equal to the initial investment I_o, plus any new investment, if new investment is limited to retained earnings only then the earnings in each period may be expressed in the following pattern:

$$E_1 = I_o r_1$$

$$E_2 = I_o r_2 + I_1^R r_2$$

$$E_3 = I_o r_3 + I_1^R r_3 + I_2^R r_3$$

and so on. Therefore,

$$E_j = I_o r_j + \sum_{i=2}^{j} I_{i-1}^R r_j.$$

Also, since

$D_j = E_j - I_j^R$, it follows that

$$D_j = \left[I_o r_j + \sum_{i=2}^{j} I_{i-1}^R r_j \right] - I_j^R$$

Thus the general model of stock valuation under the "dividends only" approach using the rate of return and assets available to the firm may be expressed as follows:

$$V_o = \sum_{j=1}^{\infty} \left[(I_o r_j + \sum_{i=2}^{j} I_{i-1}^R r_j)(1+k)^{-j} \right] - \sum_{j=1}^{\infty} \left[I_j^R (1+k)^{-j} \right]$$

The "dividends only with finite horizon" approach again assumes that the investor intends to sell the stock at some future time. This hypothesis requires that the present value of the investment be determined by discounting the anticipated dividend stream plus the final anticipated price. In this case,

$$V_o = \sum_{j=1}^{\infty}\left[(I_o r_j + \sum_{i=2}^{j} I_{i-1}^{R} r_j)(1+k)^{-j}\right] - \sum_{j=1}^{\infty}\left[I_j^R (1+k)^{-j}\right]$$

$$+ P_n(1+k)^{-n} .$$

At this point the general model may be expanded to include the fact that growth in earnings may also occur as a result of issues of new stock. If the only manner of generating earnings in period j were the sale of new stock, then the earnings in period j would be

$$E_j = I_o r_j + \sum_{i=2}^{j} I_{i-1}^{S}(r_j - k_j')$$

where I_j^S is the new investment through sale of stock and k_j' is the earnings rate of discount acquired by purchasers on the new stock. Modifying the "dividend only with finite horizon" model to include the increase in earnings through the sale of new stock will express the value of the stock as:

$$V_o = \sum_{j=1}^{n}\left[\left(I_o r_j + \sum_{i=2}^{j} I_{i-1}^{R} r_j + \sum_{i=2}^{j} I_{i-1}^{S}(r_j - k_j')\right)(1+k)^{-j}\right]$$

$$- \sum_{j=1}^{n}\left[I_j^R (1+k)^{-j}\right] + P_n (1+K)^{-n}$$

Growth in earnings may also accrue to the stockholder when a when a firm issues bonds, as long as the firm is able

to earn a higher rate than the interest rate it is required to pay on its debt. The earnings in period j due to debt financing only may be expressed as

$$E_j = I_o r_j + \sum_{i=1}^{j} I_{i-1}^{D} (r_j - i_j) \quad .$$

Recalling the relation between earnings and dividends, the growth in earnings through outside debt financing may be incorporated into the "dividends only with finite horizon" model, including the sale of new stock. Then the value of a share of stock may be modelled as follows:

$$V_o = \sum_{j=1}^{n} \Bigg[\Big(I_o r_j + \sum_{i=2}^{j} I_{i-1}^{R} r_j + \sum_{i=2}^{j} I_{i-1}^{S} (r_j - k_j') + \sum_{i=1}^{j} I_{i-1}^{D} (r_j - i_j) \Big) (1+K)^{-j} \Bigg]$$

$$- \sum_{j=1}^{n} \Big[I_j^{R} (1+K)^{-j} \Big] + P_h (1+k)^{-h}$$

By changing earnings to reflect net cash flows and allowing for the reinvestment by prior stockholders as well as investment by new stockholders the general model, which now accommodates all five approaches to discounting and all four methods of growth, can be formulated as follows:

$$V_0 = \sum_{j=1}^{n} \Bigg[\Bigg(I_0 r_j + \sum_{i=2}^{j} I_{i-1}^{R} r_j + \sum_{i=2}^{j} I_{i-1}^{S^N} (r_j - K_j') + \sum_{i=2}^{j} I_{i-1}^{S^0} r_j + \sum_{i=1}^{j} I_{i-1}^{D} (r_j - i_j) \Bigg) (1+K)^{-j} \Bigg] - \sum_{j=1}^{n} \Big[(I_j^R + I_j^{S^0})(1+K)^{-j} \Big] + P_n (1+K)^{-n}$$

The final modification to this general model by Sloane and Reisman is to assign weighting factors to various sections of this model to allow the user to express the model in accordance with his own inclinations. To allow this, the following terminology is used:

α_E = weighting factor for the earnings/dividend term

α_I = weighting factor for the investment term

α_P = weighting factor for the terminating sale term

β = weighting factor for the term in the investment section for purchases of new stock by the initial investor

The general model in its final form, then, expresses the value of a share of stock as follows:

$$V_0 = \alpha_E \sum_{j=1}^{n} \Bigg[\Big(I_0 r_j + \sum_{i=2}^{j} I_{i-1}^{R} r_j + \sum_{i=2}^{j} I_{i-1}^{S^N} (r_j - K_j') + \sum_{i=2}^{j} I_{i-1}^{S^0} r_j + \sum_{i=1}^{j} I_{i-1}^{D} (r_j - i_j) \Big) (1+K)^{-j} \Bigg]$$

$$- \alpha_I \sum_{j=1}^{n} \left[(I_j^R + \beta I_j^{S^0}) (1+K)^{-j} \right] + \alpha_P \left[P_n (1+K)^{-n} \right]$$

Now, for example, a user could use the general model to reflect the "earnings only" hypothesis by setting $\alpha_E = 1$, $\alpha_I = \alpha_P = 0$, and $n = \infty$. Or, the "dividends only with finite horizon" approach may be modelled by setting $\alpha_E = \alpha_I = \alpha_P = 1$, and $\beta = 0$.

To greatly simplify the general model of Sloane and Reisman, it may be generalized that the present value of a share of stock is a function of earnings, rate of return, and the market price of a share of stock at the end of the horizon. Since use of the model depends upon future earnings, an uncertain rate of return, and the market price of the stock at some future time, the model will be successful in its evaluation only insofar as accurate predictions may be made of future earnings, rates of return, and prices.

Therefore, attention will now be focused on models which attempt to predict or at least give some information regarding these variables.

III. EARNINGS ESTIMATION MODELS

Evidently due to the fact that the estimation of future dividends involves even more uncertainty than is inherent in the estimation of future earnings, most of the literature in this area is devoted to earnings estimation. Elton and Gruber (Ref. 4) have done interesting work in comparing various techniques for forecasting earnings per share. They examined the accuracy of forecasts produced by nine mechanical forecasting techniques and rated their relative performances.

The nine mechanical forecasting techniques were broken down into four main categories:

1. exponentially weighted moving averages
2. regressions
3. moving averages
4. a naive model

The exponentially weighted moving averages were further categorized as follows:

1. Multiplicative exponential with no trend in trend -- permits growth in earnings per share by letting earnings grow by a constant percent each year.
2. Additive exponential with no trend in trend -- permits growth in earnings per share by letting earnings grow by a constant amount each year.

3. Multiplicative exponential with trend in trend -- permits growth in earnings per share as a constant percentage and also permits variation in growth rate.
4. Additive exponential with trend in trend -- permits growth in earnings per share as a constant amount and also permits variation in growth rate.

Two types of regression forecasts were used. The first was a linear least-squares regression and the second was a log-linear regression which permits growth to be multiplicative rather than additive.

The simple moving average forecasts were also of two types. The first type was a fixed period moving average using four periods. The second was a simple moving average in which a search algorithm was used to determine the optimum number of periods which corresponded to a weighting factor used in a similar exponentially weighted moving average.

Finally, the naive model estimated earnings by simply assuming that they were equal to earnings in the previous period plus the prior change which had occurred in earnings. Next, Elton and Gruber compared the relative performance of these nine mechanical techniques by using each of the nine models to estimate earnings for the period 1962 - 1967 on a stratified random sample of 180 firms. To test the generality of the results from the 180 firm sample, the 180 firms were further divided into three random samples of sixty firms each.

Using these samples three different types of forecasts were made:

1. One-year forecasts for each of the years 1962-1967
2. Two-year forecasts for the years 1963 and 1966
3. Three-year forecasts for the years 1964 and 1967

The statistic used to judge the accuracy of the forecasts was the square of the forecast error for each of the models. The first test used was the Friedman two-way analysis of variance by ranks, to determine if the squared errors all came from the same population. This hypothesis was rejected at the .001 level for all of the forecasts, whether of one, two, or three year duration.

Next, in order to measure differences in performance, an estimate was made of the frequency function of the differences in squared error between each pair of forecasting techniques. By application of the Central Limit Theorem, the normal distribution was used to determine whether the mean of a given frequency function was significantly different from zero. This, in turn, was used to determine the dominance of one technique over another.

As a result of their investigation, Elton and Gruber determined that among the one-year forecasts the additive exponential with no trend in trend outperformed every other technique at at least a 0.10 level of significance, and the rest of the exponential models tended to outperform the rest of the models. Next to the exponential models, the naive model and the optimum moving average model were superior to

both of the regression models and the fixed-period moving average model. For both the two-year and the three-year forecasts, the exponential models dominate every other technique at the .01 level of significance. For this length of forecast the multiplicative exponential with no trend in trend tended to outperform the other exponential models, although its dominance was statistically significant at the .10 level in only half the cases.

Another method of investigating earnings was studied by Brown and Niederhoffer (Ref. 5) who compared the forecasting ability of predictors based on annual earnings data as opposed to predictors based on quarterly earnings data. The basis for their investigation was to determine if quarterly earnings statements do in fact aid security analysts in predicting annual earnings, and whether or not interim predictors tend to perform better as the end of the financial year approaches.

The predictive models which used annual earnings data were formulated as follows:

1. A1 -- earnings for last year.
2. A2 -- earnings for last year plus the change in earnings over the most recent year.
3. A3 -- earnings for last year plus the same percentage change as occurred in earnings over the most recent year.
4. A4 -- earnings for last year plus the average change over the available history of the data.

The predictive models using quarterly earnings data were set up as follows:

1. Q1 -- four times the average quarterly earnings up to the end of the current quarter.
2. Q2 -- earnings for last year plus the same percentage change as occurred between earnings up to the current quarter of the year being examined and the same quarter of the previous year.
3. Q3 -- the sum of earnings over the most recent four quarters.
4. Q4 -- earnings from last year plus four times the difference between the average quarterly earnings up to the end of the current quarter and the same quarter the previous year.

Data was collected for each of the eight predictive models for a total of 519 firms and predictions were made by each model at the end of quarters one, two, and three for the years 1963, 1964, and 1965. Three measures of predictive accuracy were then used to determine the performance of the models. The first measure was the average percentage error of forecast. This measure was looked at in three different ways:

1. arithmetic percentage error of each model averaged over all predictions.
2. absolute percentage error of each model averaged over all predictions.
3. squared percentage error of each model averaged over all predictions.

A second measure of predictive accuracy was obtained by calculating the average ranks of the models. This was done by ranking the absolute values of the prediction errors for each firm for each of the eight models. The average rank was then computed over the sample of 519 firms. The third manner of evaluating predictive accuracy was to describe for each model its complete empirical distribution of percentage forecast error by computing cumulative relative-frequency tables of the absolute percentage errors.

Although no measure of statistical significance was given for their results, Brown and Niederhoffer came to the following conclusions:

1. On the basis of predictions made at the end of the first quarter, the quarterly predictors were generally superior to the annual models when both types were considered as a group.
2. The best of the quarterly models was consistently better than the best of the annual models.
3. As data from the second and third quarters were added, further improvement was made by the quarterly models.

IV. RATE OF RETURN ESTIMATION MODELS

The second variable which must be estimated in order to allow successful application of the general stock valuation model is the rate of return. Soldofsky and Biderman (Ref. 6) examined methods of calculating rates of return for broad groups of stocks. Specifically, they used three models to examine the yields, or rates of return, for the Dow-Jones Industrial Average (DJIA) for the period 1930-1966.

The first model considered was the internal rate of return model. This model expresses the initial market price of a share of stock as the flow of dividend payments plus the market price of a share at the end of the period discounted over the entire period. In other words, the model uses a "dividends only with finite horizon" approach, and is formulated as follows:

$$P_o = \sum_{t=1}^{n} \frac{D_t}{(1+k)^n} + \frac{P_n}{(1+k)^n}$$

In this case, P_o is the closing DJIA for the initial year, D_t represents the DJIA dividend for a given year, and P_n is the closing DJIA for the terminal year being considered. The internal rate of return, k, is then calculated using this model by an iterative technique which equates P_o with the corresponding polynomial and solves for the value of k which satisfies the equation.

The second model considered was the arithmetic mean of annual rates of return. The annual rate of return is calculated as follows:

$$r_t = \frac{D_t + P_t}{P_{t-1}} - 1$$, where P_t is the DJIA for the close of year t and D_t is the DJIA dividend for year t. The arithmetic mean of these annual rates of return is then calculated for the number of years under consideration and is used as the estimate for rate of return.

The third model used was the geometric mean of annual rates of return. The annual rates of return were calculated as before and the geometric mean was expressed as follows:

$$\bar{r} = \sqrt[n]{\prod_{t=1}^{n} (1 + r_t)} - 1$$

This particular formulation is used because the geometric mean cannot be used with negative numbers, so unity is added to each rate of return, and after the root is taken unity is subtracted.

Soldofsky and Biderman used these three models to compute the rates of return for the DJIA for the period of 1930 through each succeeding terminal year up to and including 1966. The differences in the rates of return for some of the years are particularly large, although no measure of statistical signficance was given. Soldofsky and Biderman seem to

prefer the use of the geometric mean of annual yields model to the other two models as the internal rate of return model assumes that the return is reinvested each year at the average rate of return for the entire period, and the arithmetic mean model gives extreme values a heavier weight than may be desired.

Another manner of investigating rates of return is to assume a particular model of rate of return to apply and then to construct a model using explanatory variables to determine how those variables affect the rate of return in different years. An example of this type of investigation is the work of Nerlove (Ref. 7). Nerlove defines rate of return as the value of r that equates the discounted value of the flow of dividends during a period of years and the price at the end of the period to the initial price. Thus, the rate of return is calculated from the following formulation:

$$P_o = \sum_{t=0}^{T-1} \frac{d_t}{(1+r)^{t+1}} + \frac{P_t + d_t}{(1+r)^{T+1}}$$

where P_o is the initial price of a share of stock, T is the terminal year of the investigative horizon, and d_t is the dividend per share paid in year t. Price, whether P_o or P_T, was interpreted as an arithmetic average of the high and low prices for the year in question. The internal rate of return r is then calculated by an iterative technique which finds the value of r that satisfies the equation.

In his investigation Nerlove calculated the rates of return for 371 firms for three five-year sub-periods; 1950 - 1954, 1955 - 1959, and 1960 - 1964, as well as for the entire period 1950 - 1964. The mean of the 371 observations in each sub-period was calculated and used as the observed value for rate of return for that sub-period. Then Nerlove constructed two multiple linear regression models which he used to estimate the rate of return for each of the 371 firms during each of the four periods mentioned above. The first model used the following eight explanatory variables plus the constant term:

1. Rate of growth of net sales -- computed for each firm by taking the regression of net sales on a simple linear trend for the period under consideration and dividing the slope of time in the regression by the mean net sales for the period.
2. Rate of growth of earnings available for common -- computed for each firm in a similar manner to the rate of growth of net sales.
3. Mean retained earnings per dollar of total assets -- computed for each firm for each year by deducting common dividends from earnings available for common and dividing by total assets. The value used in the regressions was the arithmetic mean of the yearly value over the period being investigated.
4. Mean dividends per dollar of total assets -- computed for each firm by dividing yearly dividends paid on

common stock by yearly total assets. The arithmetic mean of these yearly values for the period in question was used in the regressions.

5. Mean reciprocal of leverage -- computed for each firm on a yearly basis by dividing equity values by the total of long term debt, preferred stock, and common equity. The arithmetic mean of these values for the period was used in the regression.
6. Mean inventory turnover -- the value used in the regression was the yearly ratio of cost of goods sold to inventory value for each firm, arithmetically averaged over the particular period.
7. Mean share turnover -- computed for each firm as the ratio of the number of common shares traded during a year to the number of common shares outstanding at years end. The arithmetic mean of these ratios over the particular period was used in the regressions.
8. Mean gross plant per dollar of total assets -- computed for each firm as the value of the plant divided by total assets and averaged over the period in question.

The regression of rate of return on the eight explanatory variables was run using the values determined for the 371 firms. Using the observed value of the rate of return as previously described, the multiple correlation coefficient R^2 was determined for each of the four periods, as well as the coefficients of each of the explanatory variables and

their t-ratios. The basic results of the first regression model may be summarized as follows:

1. The most important variables over all periods both short and long were the rate of growth of sales and the retained earnings per dollar of assets.
2. The rate of growth of earnings was also a significant variable in all periods, but appeared to be of lesser importance than sales growth and retention of earnings.
3. The effect of the other variables was inconclusive, but of little importance in comparison to the other three.
4. The highest value of R^2 achieved was .515 for the period 1955 - 1959.

A second regression model was formulated by Nerlove which used the same eight explanatory variables as used in the first regression model, but added a total of 23 additional variables. Twenty-one of the additional variables were dummy variables used to reflect classification of the 371 firms into 21 industry related groupings, and the other two variables were dummy variables used to represent firms which were significantly conglomerate or misclassified. The results of this regression showed that the addition of the dummy caused a substantial increase in the R^2 values, from .515 to .587 for the 1955 - 1959 period, for example. Furthermore, the extremely large t-ratios for many of the industry dummies indicated that many of them were individually significant.

Finally, Nerlove examined the stability of the sign, significance, and size of the coefficients of the total set of explanatory variables by performing regressions over periods beginning with a length of three years, and increasing in steps of one year up to a fifteen year period. This process determined, among other things, that sales growth and earnings retention are positively associated with the rate of return and were highly significant whether or not the industry dummies were used.

Another method of looking at rate of return was proposed by Benishay (Ref. 8). He defined the measured rate of return for an equity as the weighted average of annual earnings for nine consecutive years divided by the market value of that equity in the ninth year. The earnings were weighted such that 10/18 of the weight was assigned to the ninth year and 1/18 of the weight was assigned to each of the preceding eight years. The market value used was the arithmetic mean of the high and low values of the equity in the ninth year. This formulation was used to compute the measured rate of return for a total of 56 firms for the years 1954 - 1957, with a value calculated for each firm in each of the four years.

Using these values as the observed values for rate of return, Benishay then proposed a cross-sectional multiple linear regression model to determine a measured rate of return for each of the 56 firms. The following independent variables were used in the model:

1. X_1 - Growth in earnings - computed by running a simple linear regression of after-tax earnings for each firm on time for a period of nine years, including the cross-section year and the eight preceding years. The earnings coefficient was divided by the arithmetic mean earnings for the same period and this value was used in the regression.
2. X_2 - Growth in equity value - computed for each firm by running a simple linear regression of high and low equity values on time for the nine-year period. The coefficient of equity value was divided by the arithmetic mean equity value for the same period of time and this value was used in the regression.
3. X_3 - The pay-out ratio - computed as the value of dividends paid times 100 divided by earnings for each firm for the cross-section year and the two preceding years. The log of the arithmetic mean of these values was used in the regression.
4. X_4 - Stability of income - computed by forming a ratio whose numerator was the arithmetic mean of after-tax earnings for the cross-section year and the eight preceding years. The denominator was the standard deviation of these same observations. The log of this ratio was used in the regression.
5. X_5 - Stability of equity value - computed by forming a ratio whose numerator was the arithmetic mean of the high and low market values of the firm's equity for the

cross-section year and the eight preceding years. The denominator was the standard deviation of these same observations. The log of this ratio was used in the regression.

6. X_6 - Size - computed as the log of the arithmetic mean of the high and low values of the firm's equity in the cross-section year.
7. X_7 - Debt-equity ratio - computed as the book value of debt at the end of the cross-section year divided by the value for variable X_6.

The first three independent variables were intended as corrective variables in that it was hoped that they would tend to remove errors obstructing a valid measurement of the theoretical concept of a rate of return on equity capital. The other four independent variables attempt to measure the differential "desireability" of holding equities.

Several cross-sectional multiple regressions were run for each of the four years 1954 through 1957, using values computed for the 56 firms. Some regressions were run using all of the independent variables and others were run in which some of the variables were excluded. The following results were deemed noteworthy by Benishay:

1. Multicollinearity existed between the two growth variables X_1 and X_2. Regressions conducted when one of the variables was excluded and the other included showed that the growth in equity value, variable X_2, performed better as a corrector than did the earning trend, variable X_1.

2. The size variable, X_3, was the most consistent and most significant statistically over the four regressions. It had a negative relation with the rate of return in all four cases, which indicated that, all other things being equal, the market tended to prefer larger to smaller firms.
3. The coefficients for stability of equity, variable X_5, had positive signs in all four regressions, which indicated that equity stability was avoided.
4. The coefficients for stability of income, variable X_4, had negative signs in all four regressions, which indicated a preference for stability of earnings.

V. PRICE ESTIMATION MODELS

A. THE RANDOM-WALK HYPOTHESIS

A major stumbling block toward any practical application of the general stock valuation model is the estimation of the price of a given share of stock at any future time. One of the difficulties involved in this estimation is that there is a great deal of disagreement on the general behavior and distribution of stock prices. As Mandelbrot and Taylor (Ref. 9) reported,

> "There are at least four schools of thought on the statistical distribution of stock price differences, or more generally, stochastic models for sequences of stock prices. In terms of numbers of followers, by far the most popular approach is that of the so-called 'technical analyst', phrased in terms of short-term trends, support and resistance levels, technical rebounds, and so on. Rejecting this technical viewpoint, two other schools agree that sequences of prices describe a random walk, where price changes are statistically independent of previous price history, but these schools disagree in their choice of the appropriate probability distributions and/or in their choice of the appropriate 'time' parameter (the physical time - days, hours - or a randomized operational time ruled by the flow of transactions). Some authors find the time changes to be normal or Gaussian, while the other group find them to follow a stable Paretian Law with infinite variance. Finally, a fourth group (overlapping with the preceding two) admits the random walk as a first-order approximation but notes recognizable second-order effects."

The basic idea of the random-walk hypothesis may be gleaned from an examination of some definitions proposed by various authors. For example, Smidt (Ref. 10) said,

> "The basic idea behind the random-walk hypothesis is that in a free competetive market the price currently quoted for a particular good or service should reflect all of the information available to participants in the market

that influence its present price. To the extent that future conditions of the demand or supply are currently known, their effect on the current price should be properly taken into account.

A statistical process which has the property that the expected future value of a random variable is independent of past values of the variable is said to be a Martingale process. What is generally referred to as the random-walk hypothesis requires that in a perfectly competetive market, price changes should be the outcomes of a Martingale process."

Another interpretation was given by Seelenfreund, Parker, and Van Horne (Ref. 11) who said,

"This theory implies that past stock-price movements cannot be used to predict future market prices in such a way as to 'profit' from the predictions. By not profiting we mean that a trader using the past history of stock prices cannot apply mechanical decision rules that result in a consistently better performance than a simple buy-and-hold strategy. If stock-price movements were to become systematic so that a 'profit' were possible, proponents of the random-walk theory argue that a sufficient number of market participants would quickly recognize the recurring pattern and exploit it. In exploiting it they would drive out the opportunity for 'profit' causing the price series to approximate a random walk.

Thus, the random walk theory presumes that information is freely and readily available and that there are enough market participants with sufficient resources to take advantage of any profit opportunity arising from systematic price movements of an individual stock. These participants compete against each other making all non-random fluctuations so small that they cannot be exploited profitably."

As a final example, Jones and Litzenberger (Ref. 12) said,

"Random-walk theorists deny the existence of intermediate stock price trends. They maintain that the stock market is a highly competetive or efficient market. In such a market, new information would be available to a large number of market participants. They argue that price adjustments would be instantaneously established by the competetive interaction of these knowledgeable and rational investors. Common stock prices would reflect intrinsic values and

published financial data would be irrelevant to investment selection. The best investment strategy in such a market would be random selection."

Just as there is a wide range of definitions of the random-walk hypothesis, there are also many theoretical models used to describe it in the literature. For example, Granger (Ref. 13) investigated the properties of the following model:

$$P_t = P_{t-1} + \varepsilon_t$$

where P_t represented the price series and ε_t was a purely random, white noise series such that ε_t and ε_s were uncorrelated for all $t \neq s$.

Another model was proposed by Osborne (Ref. 14), who defined a random variable

$$Y = \ln \left[\frac{P(t+\tau)}{P_o(t)} \right]$$

where $P_o(t)$ and $P(t + \tau)$ were the price of the same random choice stock at random times t and $t + \tau$. Osborne's hypothesis was that the steady state distribution function of Y was

$$\phi(Y) = \exp\left(\frac{-Y^2}{2\sigma^2\tau}\right) \Big/ \sqrt{2\pi\sigma^2\tau}$$

where σ was the dispersion developed at the end of unit time.

Still another model was formulated by Press (Ref. 15) as

$$Z(t) = C + \sum_{k=1}^{N(t)} Y_k + X(t)$$

which uses the following notation:

1. Z(t) is the natural logarithm of the price of a given security at time t.
2. C = Z(0) and is assumed to be known
3. $(Y_1, Y_2, \ldots, Y_k, \ldots)$ is a sequence of independent identically distributed $N(\theta, \sigma_2^2)$ random variables.
4. N(t) is a Poisson process with parameter λt, which represents the number of random events which cause price changes occurring in time t. N(t) is independent of the Y_k.
5. The process $\{X(t), t \geq 0\}$ is a Wiener process independent of both N(t) and the sequence of Y_k variables. X(t) is distributed $N(0, \sigma_1^2)$.

B. PRICE MODELS

Most of the models dealing with prices and price estimation were formulated to test various facets of the random-walk hypothesis, and especially its major assumption, that of independence. One method of testing involves the use of standard statistical tools. There are numerous articles in the literature relating to tests of this type; some of which support the random-walk theory and others which do not. An example of this type of testing is found in the work of Kemp and Reid (Ref. 16), who studied the British stock market.

Kemp and Reid used as their data base the daily closing prices of individual common stocks as listed in the Financial Times for the period of 28 October 1968 to 10 January 1969, giving 52 observations for each firm studied. The 50 firms

chosen for study were selected on the basis of proportional stratified random sampling. The strata were selected on the basis of industrial groupings used by the Financial Times. The data obtained for the sample then consisted of 52 daily closing prices for each of 50 firms, plus the same observations for the Financial Times' Index. The data were adjusted for dividends and stock splits.

Rather than become involved in the controversy over the distribution of stock prices, Kemp and Reid opted to use a series of non-parametric tests to test for randomness in the sample. The first test used was the run test, which used the median as a dividing point to determine the different runs. This test detected six cases in which the null hypothesis of randomness was rejected at the .05 level of significance.

The second test used was the Wallis-Moore test for cycles. This test is based on the number of runs of like signs in the first differences of the price series. The test statistic, which is similar to the chi-square test statistic, is computed as

$$\chi^2_P = \frac{(u_1 + U_1)^2}{U_1} + \frac{(u_2 + U_2)^2}{U_2} + \frac{(u_3 + U_3)^2}{U_3}$$

where the u_i are the actual number of runs of length one, two, and three or more, and the U_i are the number of runs of length one, two, and three or more that would be expected from a random sample. This test detected 24 cases in which the null hypothesis of randomness was rejected at the .05 level

of significance. Due to the fact that in many of the samples several days might pass with no change in the daily closing prices, zeros were present in the first differences. Because of this, the sample size was reduced to eliminate the zeros and the Wallis-Moore test was recomputed for the reduced samples. In this case there were nine cases where the null hypothesis was rejected at the .05 level.

The third test used was based on the total number of runs in the first differences. In this case the test statistic, the total number of runs, is approximately normally distributed with mean (2n - 7) / 3 and variance (16n - 29) / 90, provided n, the sample size, is greater than ten. This test yielded a total of 36 cases in which the null hypothesis was rejected at the .05 level. This test was also recomputed, based on a smaller sample size, after the zeros had been discarded. In this case there were 23 cases of rejection of the null hypothesis at the .05 level.

The fourth test used was the Wald-Wolfowitz test for auto correlation, which was based on the test statistic

$$R = \sum_{t=1}^{N-1} \Delta P_t \Delta P_{t+1} + \Delta P_n \Delta P_1$$

where P_i is the closing price for the ith trading day. This test produced twelve cases where the null hypothesis was rejected at the .05 level.

By aggregating their results, Kemp and Reid found that approximately eighty percent of the sample showed significant

non-randomness by at least one of the tests employed, although this figure was reduced to approximately fifty percent when the zeros were discarded from the first difference series and the tests were recomputed using the smaller sample size. This was viewed as an indication that the random-walk hypothesis had been over-generalized. However, the observations based on the daily closing prices of the Financial Times Index were not found to be non-random by any of the tests employed. This same result had been found in other investigations of this sort run on aggregate data in the form of an index.

In addition to the use of standard statistical tests, other methods have been used to test the validity of the random-walk theory which formulated mechanical trading rules based on the price series of either individual stocks or aggregate groups of stocks. These rules were then used to simulate trading over a period of time and at the end of the period the profits gained from application of the trading rules were compared to the profits which would have been generated from pursuing a simple buy-and-hold strategy. Results which indicated a greater profitability under the mechanical trading rules would have tended to refute the random-walk theory. The rationale behind this sort of testing was explained by Fama and Blume (Ref. 17) who said,

> "The basic hypothesis of the (random-walk) theory is that successive price changes in individual securities are independent random variables. Independence implies, of course, that the past history of a series of changes cannot be used to predict future changes in any 'meaningful' way.

What constitutes a 'meaningful' prediction depends, of course, on the purpose for which the data are being examined. For example, the investor wants to know whether the history of prices can be used to increase expected gains. In a random-walk market, with either zero or positive drift, no mechanical trading rule applied to an individual security would consistently outperform a policy of simply buying and holding the security. Thus the investor who must choose between the random-walk model and a more complicated model which assumes the existence of an excessive degree of either persistence (positive dependence) or reaction (negative dependence) in successive price changes, should accept the theory of random walks as the better model if the actual degree of independence cannot be used to produce greater expected profits than a buy-and-hold policy.

On the other hand, the statistician has different though equally pragmatic conditions of what constitutes an important violation of the independence assumption of the random-walk model. He will typically be interested in whether the degree of dependence in successive changes is sufficient to account for some particular property of the distribution of price changes or whether the dependence is sufficient to invalidate the results produced by statistical tools applied to the data. For example, price changes may be one variable in a regression analysis and the statistician will want to determine whether dependence in the series might produce serial dependence in the residuals. If the amount of dependence is low, he will probably conclude that it will not seriously damage his results. From the investor's point of view, however, the dependence may make the expected profits from some mechanical trading rule greater than those of a simple buy-and-hold policy.

It is important to note, however, that though a strict definition of 'important dependence' is always specific to the case at hand, the ultimate criterion is always practical. In an encounter with a more complicated alternative, the theory of random walks is overthrown only if the alternative leads to a better action than the random-walk theory would have suggested."

One example of this type of test, reported by James (Ref. 18) used monthly moving averages, coupled with threshhold constraints around the moving average line, to generate trading signals. Two basic models were used. The first was an unweighted moving average of the form

$$W_g = \frac{\sum_{t=g-N+1}^{g} \frac{X_t}{N}}{N}$$

where W_g was the moving average at time g, X_t was the price at time t, and N was the number of observations. This simple moving average was constructed for an observational period of seven months, so in this case g = 7, 8, ..., n, where n was the terminal time period, and N = 7. For the simple moving average the threshhold constraints, called MAX and MIN, were set to values of 1.02 and .98 respectively. The trading rule developed by this model was to compare the monthly closing price for each month after the seventh month to the moving average value W_g. Then, when the closing price rose above the moving average value by two percent or more, a buy signal was generated, and if the security was not held in a long condition at the time a purchase was simulated at the current price. Conversely, if the closing price fell below the moving average by two percent or more, a sell signal was generated, and if the security was not held in a short position at the time a sale was simulated at the current price.

The second model was an exponentially-smoothed moving average, expressed as

$$F_i = F_{i-1} + \alpha(X_i - F_{i-1})$$

where F_i was the value of the moving average at period i, α was the smoothing constant, and X_i was the closing price of

the security at period i. An observational period of six months was used before values of this moving average were used to generate the trading rules described above.

The data used in the experiment came from a tape of month-end common stock price relatives covering the period 1926 - 1960. The basic experiment itself consisted of selecting a sample of stocks and observing their behavior for an extended period of time, the exact length of which was not specified. For each stock two computer simulations were run. In the first run, a simulated portfolio manager was given $100. After the initial period of observation the value of the moving average was compared to the month-end price of the stock. Based on the trading rules described above the computer simulated a buy of as many shares as could be purchased at the closing price, at the first available opportunity. This position was held until the price movement reversed itself enough to generate a reverse indicator, at which time the position was closed out, and the funds thus obtained were utilized to initiate a position in the opposite direction. In the second simulation the computer simulated a buy-and-hold procedure at that month when the first signal was given under the moving average rule.

The basic experiment consisting of the two simulation runs was repeated for both the unweighted moving average and the exponentially-smoothed moving average under conditions of allowing various combinations of long and short positions, for various sample sizes ranging from 232 to 1376 stocks,

and with various values of the smoothing coefficient, α, ranging from 0.1 to 0.4. In addition, the experiments were repeated using two different types of price data, one which included dividends and one which did not.

At the end of each experiment two values were calculated, one reflecting the value of the portfolio managed according to the trading rules generated by the moving average, and the other the value of the portfolio managed by the buy-and-hold rule. These values were called XPOT and YPOT respectively. Then for each of the securities in the sample the value

$$D_i = XPOT_i - YPOT_i \quad (i = 1, 2, \ldots, n)$$

was calculated. $\bar{D}$, the sample mean, was used as an estimate of D, the true difference. The null hypothesis, H_o: $E(D) > 0$ (that the moving average trading rules are effective and will produce better results than a buy-and-hold rule) was tested against the alternative hypothesis H_1: $E(D) \leq 0$ (that the moving average trading rules are ineffective) was tested using a paired t-test and was rejected for all cases at the .01 level of significance, and for all but two cases at the .05 level. The results using the moving average trading rules were consistently inferior to the buy-and-hold rule when the effect of dividends was included. James therefore concluded that the results of his investigation lent support to the random-walk theory.

Alexander (Ref. 19) proposed another set of mechanical trading rules based on the use of a logarithmic filter applied to peak and trough prices to generate buy and sell

signals. Alexander's trading rules were generated by applying this filter to the logarithm of the daily closing price of a particular industrial average. The filter used may be referred to as an X percent filter of arithmetic size F and logarithmic size f. For example, for a one percent filter the arithmetic size F is 1.01, and the logarithmic size f is .00432, the logarithm of 1.01. In applying the filter, the first step is the choice of a filter size. Then the daily closing price of the particular average being studied is examined until a trough appears. At this point the upswing is studied. At the point where the log of the closing price of the average exceeds the log of the trough closing price by .00432 or more (assuming a one percent filter had been chosen) a buy signal is generated. On the other hand, if a peak comes first, then the downswing is studied, and a sell signal is generated at the point where the log of the peak closing price exceeds the log of the average closing price by .00432 or more.

Alexander used filters ranging in size from five percent to fifty percent applied to the following three data bases:

1. Daily closing price of the Dow-Jones Industrial Average from 2 January 1897 to 3 September 1929.
2. Daily closing price of Standard and Poor's Industrial Average from 12 December 1914 to 3 September 1929.
3. Daily closing price of Standard and Poor's Industrial Average from 7 September 1929 to 31 December 1959.

Using the trading rules generated by the specified filter as applied to the particualr industrial average being studied Alexander then calculated the following variables:

1. Average move -- for each move corresponding to a specified filter, the variable M is the difference between the logarithms of the upper and lower end points of the move. $\overline{M}$ is the arithmetic mean of the M's. The value for the average move is 100 (antilog $\overline{M}$ - 1).
2. Bias -- for a given filter the trigger price is that price whose logarithm would just generate a trading signal based on the specified filter. Since in many cases the actual price at which the stock may be bought or sold may be above or below the trigger price, a bias is introduced. The bias variable, b, represents the average difference between the trigger price and the actual closing price on the day on which the filter was triggered, as a percentage of the trigger price.
3. Average logarithmic profits -- this variable is defined as $\overline{R} = \overline{M} - (2b + 2f)$, where $\overline{M}$ is the average move, b is the bias variable, and f is the logarithmic filter size.
4. Average percent profit on an upmove -- for a set of up and down moves whose average size is $\overline{M}$ the average percent profit on an upmove is defined as

$$\overline{P}_u = 100 \left[\text{antilog } (\overline{R} + \frac{d}{2}) - 1 \right] \text{ where } \frac{d}{2} \text{ is defined as } \frac{1}{N \left[\log (\text{last price}) - \log (\text{initial price}) \right]}, \text{ with N being}$$

the total number of transactions (a transaction being a purchase and a sale).

5. Average percent profit on a downmove -- this variable is defined as

$\bar{P}_d = 100\ (1 - \frac{F^*}{g + 1})$ where $F^* = e^{f + 2b}$ and $g = f + 2b + \bar{R} - \frac{d}{2}$.

6. Average profit per transaction -- this variable is defined as

$$\bar{P} = (100 + \bar{P}_u)^{\frac{1}{2}} (100 + \bar{P}_d)^{\frac{1}{2}} - 100 .$$

7. Average profit per year -- this variable is defined as

$\bar{P}_y = 100 \left[(1 + \frac{\bar{P}}{100})\right]^{\bar{q}} - 100$, where $\bar{q}$ is the average number of transactions per year.

The results of Alexander's investigation may be interpreted in different ways. In almost every period for almost every size filter the average profit per year was less using the filter generated trading rules than the profit that would have been realized by using a simple buy-and-hold strategy over the same period of time. Furthermore, when commissions were deducted profits were further decreased, to the point that initial capital was all but wiped out for the smaller filter sizes which presented a large number of transactions. However, Alexander claimed that from a theoretical point of view commissions should not be deducted in a test of the random-walk hypothesis, and that furthermore the profits gained from the application of the filter trading rules should not be tested against a buy-and-hold procedure but rather

against the profits that should be expected if the random-walk hypothesis were valid. Using this criterion the filter did generate higher profits than the hypothetical expected profits in most cases when commissions were ignored, thus indicating to Alexander that there was some doubt as to the validity of the random-walk hypothesis.

Fama and Blume (Ref. 17) claimed that Alexander's results were difficult to interpret because the effects of dividends cannot be used to adjust commonly used price indices like the Dow-Jones Industrial Average and others. Therefore, they applied Alexander's filter technique to series of daily closing prices for each of the individual securities that make up the Dow-Jones Industrial Average. Using approximately a six to seven year observational period, they obtained thirty samples with from 1,200 to 1,700 observations per sample. Twenty-four different filters ranging in size from one-half a percent to fifty percent were applied to each sample. For each simulation, that is a run over the observational period using a given filter size for a given security, the reference price, whether it were a peak or a trough, was adjusted on ex-dividend days by adding back the amount of the dividend. Also, if a stock split occurred while a particular position was open, either long or short, the price of the stock before the split was adjusted upward by the appropriate factor until the position was closed.

As previously noted, Fama and Blume feel that from a practical point of view the "profitability" of any mechanical

trading rule must be compared against a buy-and-hold strategy in order to test the validity of the random-walk hypothesis. In this investigation, profitability was measured for each security and filter size by calculating a nominal annual rate or return under both the filter-generated trading rules and for a buy-and-hold strategy. However, the buy-and-hold rate of return was computed only for the period during which active positions were open under the filter rules, rather than for the entire period. The results of this investigation showed that even when broker's commissions were ignored only two securities out of thirty were more profitable under the filter technique than under the buy-and-hold strategy. When commissions were included the filter rule failed to outperform the buy-and-hold technique in all thirty securities. In fact, only four of the securities showed positive average returns per filter when the commissions were included.

Another model used to formulate mechanical trading rules was devised by Seelenfreund, Parker, and Van Horne (Ref. 11). They viewed the currently observed price of a security as the best estimate of the security's intrinsic value and interpreted the random-walk hypothesis as producing a stock price series with random fluctuations about the intrinsic value.

In an attempt to detect changes in the intrinsic value of a security, they fitted a second-degree polynomial sequentially in time to a particular price series. The coefficients of the polynomial were then updated daily in the following manner:

If P_τ is the price observed at time τ ; a_τ , b_τ , and c_τ are the coefficients of the quadratic computed at time τ ; and $\overline{P}_\tau, t = a_\tau + b_\tau t + \frac{1}{2}c_\tau t^2$ is the price at time $\tau + t$ predicted at time τ , then if $a_{\tau-1}$, $b_{\tau-1}$, $c_{\tau-1}$, and P_τ are known, the coefficients at time τ are updated to be

$$a_\tau = P_\tau + (1-\alpha)^3 \left[\overline{P}_{\tau-1,\,1} - P_\tau\right]$$

$$b_\tau = b_{\tau-1} + (c_\tau - 1)\left(\frac{3\alpha^2}{2}\right)(2-\alpha) \left[\overline{P}_{\tau-1,\,1} - P_\tau\right]$$

$$c_\tau = c_{\tau-1} - \alpha^3 \left[\overline{P}_{\tau-1,\,1} - P_\tau\right]$$

where $\overline{P}_{\tau-1,\,1} = a_{\tau-1} + b_{\tau-1} + \frac{1}{2}c_{\tau-1}$, and α is the smoothing factor. The coefficient a_τ is the expected price at time τ , b_τ is the trend in price, and c_τ is the change in trend as estimated at time τ .

Besides updating the coefficients of the quadratic at each point in time the mean-absolute-deviation of the sum of the forecast errors was estimated for one period ahead according to the formulation

$$MAD_\tau = (1 - \alpha)\, MAD_{\tau-1} + \alpha\left|\overline{P}_{\tau-1,\,1} - P_\tau\right| .$$

Then, assuming that the forecast errors were normally distributed with mean zero, a ninety-five percent confidence interval was constructed based on the finding that there is approximately a five percent probability that the absolute value of the cumulative sum of the forecast errors will exceed

$\sqrt{\frac{\pi(2-\alpha)}{1-(1-\alpha)^6}}$ times the mean-absolute deviation of the sum of the forecast errors.

Using these computational tools, Seelenfreund, Parker, and Van Horne generated the following trading rules:

1. If the cumulative sum of actual price minus forecasted price fell above the ninety-five percent confidence interval a buy signal was generated.
2. If this value fell below the ninety-five percent confidence interval a sell signal was generated.

This model was tested on a random sample of thirty industrial stocks listed on the New York Stock Exchange over a period from 1 January 1960 to 30 June 1966, and was based on the closing price of each stock for each trading day during the period. A 200 day initialization period was allowed to estimate the coefficients for the predictive model, so all tests actually began on 17 October 1960. All prices were adjusted for stock splits and dividends, and the model was tested for three different values of the smoothing factor, namely, $\alpha = .01$, $\alpha = .05$, and $\alpha = .10$. Each of these variations was tested utilizing a long position only, in which a purchase was simulated whenever a buy signal was generated and the position was liquidated when a sell signal was given. Additionally, each variation was tested utilizing both long and short positions, by simulating the taking of a short

position each time a long position was liquidated. The following additional rules were used in completing the simulation:

1. Equal amounts of money were allocated for investment in each of the thirty stocks and these amounts were segregated so that funds from the sale of one stock could not be used to purchase another.
2. Brokerage commissions of one percent were assumed for each purchase, sale, and short sale.
3. The price at which each transaction was made was the daily closing price after which a buy or sell signal was given.

For each variation of the trading rules, the average annual compound rate of return was computed both before and after brokerage commissions were taken. For the variation allowing long positions only, this average return included both capital gains or losses over the holding periods and any dividends paid while the stock was held. In the variation in which both long and short positions were allowed, it was assumed that the trader reimbursed the lender for any dividends paid while maintaining a short position, so these dividends were subtracted when calculating the average return.

Without stating any levels of significance, Seelenfreund, Parker, and Van Horne found that in the majority of cases, a buy-and-hold strategy resulted in an average annual return greater than that gained using the trading rules, after

brokerage commissions were deducted. In analyzing the different smoothing factors it was found that the trading rules worked better with the smallest smoothing factor, $\alpha = .01$. Additionally, it was found that the companies whose stocks performed better under the trading rules were consistently the poorest performers over the period examined, so that the performance of the trading rule could largely be attributed to either holding cash or selling short for a security whose price was on a downward trend during the observation period. Thus, it was concluded that the results of this investigation were consistent with the random-walk theory.

Other models have been proposed to test relationships between the price of securities and other variables, and some of these models have implications for the random-walk hypothesis even though they were not specifically designed to test it. For example, Mayor (Ref. 20) investigated the belief held by many financial analysts that there exists a positive relationship between the level of short interest in equities, defined as the number of shares borrowed, sold, or still uncovered or owed at any given time, and future movements in the prices of these equities. This belief appeared to be based on the assumption that as short traders attempt to cover their short positions they will push the price of the security upward at some future time.

Mayor investigated this hypothesis in two ways; first by means of a multivariate regression, and secondly by comparing short trading activities with a simulation model based on a

random timing of short interest positions. In the multivariate analysis four different regression models were used:

1. Model I

$$P_t = \alpha_1 + \beta_1 SI_{t-.5} + U_{1t}$$

where P_t is the price in month t, $SI_{t-.5}$ is the short interest variable lagged two weeks, and U_{1t} is a stochastic variable.

2. Model II

$$P_t = \alpha_2 + \beta_2 SI_{t-.5} + \gamma_2 t + U_{2t}$$

where $\gamma_2 t$ is a trend variable.

3. Model III

$$P_t d_t = \alpha_3 + \beta_3 \frac{SI_{t-.5}}{d_{t-.5}} + \gamma_3 t + U_{3t}$$

where d_t is an index of stock splits.

4. Model IV

In this model the error term is assumed to follow a second order autoregressive scheme, that is,

$$U_t = \rho_1 U_{t-1} + \rho_2 U_{t-2} + \varepsilon_t$$

where ε_t is a normally distributed random variable with zero mean. Thus this model assumes the form

$$P_t d_t = \alpha_4 + \rho_1 P_{t-1} d_{t-1} + \rho_2 P_{t-2} d_{t-2}$$

$$+ \beta_4 \left[\frac{SI_{t-.5}}{d_{t-.5}} - \frac{\rho_1 SI_{t-1.5}}{d_{t-1.5}} - \frac{\rho_2 SI_{t-2.5}}{d_{t-2.5}} \right]$$

$$+ \gamma_4 \left[t - \rho_1 (t-1) - \rho_2 (t-2) \right] + \varepsilon_t$$

These four models were used to make estimates from a fourteen firm random sample selected from stocks that are frequently shorted on the New York Stock Exchange, as well as an aggregate of Standard and Poor's 500 stock index. For the individual firms P_t referred to the average high and low price in the second week after the short interest was reported. For the aggregate data, P_t referred to the average of the high and low values of Standard and Poor's index in the second week after the short interest was reported. In this case the index of aggregate stock splits and dividends was based on the New York Stock Exchange estimates of total listed shares. The data used was for a four-year period ending in early 1966, providing a sample size of 48.

Mayor found that the estimates of the coefficients of the short interest variable were significant for the naive Model I, but decreased in significance as the models became less and less naive, until with Model IV the null hypothesis that $\beta_4 = 0$ could not be rejected for any of the fourteen firms at the .05 level of significance. Mayor felt that these results tended to refute the belief that short interest levels had a positive effect on future price movements, or at least suggested that the short interest level was not useful as a means of predicting future stock prices.

If this premise were accepted, Mayor reasoned that short traders as a group would tend to do no better or worse than might be expected on the basis of chance. In order to test this hypothesis Mayor proposed the following model:

$$V_i = -SI_1^i P_1^i + \sum_{t=2}^{n-1} \frac{F_t^i}{(1+r)^{t-1}}$$

$$+ \frac{\left[2AP_{n-1}^i - \frac{P_n^i d_n^i}{d_{n-1}^i}\right] SI_{n-1}^i}{(1+r)^{n-1}}$$

where

V_i = present value of short trading activities in firm i

SI_t^i = the short interest position for firm i in month t

d_t^i = index of stock distribution for firm i

n = number of periods observed

r = discount rate

Z_t^i = number of shares of firm i covered in period t, as expressed by

$$Z_t^i = \left(\frac{SI_{t-1}^i d_t^i}{d_{t-1}^i}\right) - SI_t^i$$

F_t^i = cash flow for firm i during period t, expressed as

$$F_t^i = \begin{cases} 0 & \text{when } Z_t^i = 0 \\ P_t^i Z_t^i & \text{when } Z_t^i < 0 \\ (2AP_t^i - P_t^i) Z_t^i & \text{when } Z_t^i > 0 \end{cases}$$

AP_t^i = the average price at which shares of firm i in a short position at time t were sold, expressed as

$$AP_t^i = \begin{cases} P_1^i & \text{when } t = 1 \\ \dfrac{SI_{t-1}^i AP_{t-1}^i - P_t^i Z_t^i}{SI_t^i} & \text{when } t > 1,\ Z_t^i < 0 \\ \dfrac{AP_{t-1}^i d_{t-1}^i}{d_t^i} & \text{when } t > 1,\ Z_t^i \geq 0 \end{cases}$$

These values of V_i were then compared against a simulated value V_i^* obtained by randomly selecting the actual short interest levels adjusted for splits (SI_1^i/d_1^i, SI_2^i/d_2^i,..., SI_n^i/d_n^i) such that the first number selected becomes SI_1^{i*}/d_1^i and so on. By assuming normality, the null hypothesis that V_i was significantly different from V_i^* was rejected at the .05 level for all fourteen firms, thus supporting the hypothesis that short traders do not perform significantly better than chance. Mayor felt that these results tended to lend support to the random-walk hypothesis.

Somewhat different results were found by Jones and Litzenberger (Ref. 12) in a study relating quarterly earnings reports to intermediate stock price trends. They hypothesized that changes in the estimation of the intrinsic worth of an equity on the part of stock market professionals would cause gradual price adjustments to take place over time, thus generating intermediate price trends. These gradual

adjustments would be attributed to the fact that the changes in estimation of the value of a security on the part of the market professionals would gradually be disseminated to the general investing public through advisory services, newsletters, stock brokers, etc. They further hypothesized that a quarterly earnings report that was significantly higher than that anticipated by market professionals would tend to cause upward revisions in their estimated worth of that equity and that these revisions would give rise to a gradual upward adjustment of the price of the stock.

The method used by Jones and Litzenberger to study this hypothesis was to select two samples, the first covering 510 companies for the period 1962 - 1965, and the second covering 618 companies for the period 1964 - 1967. Quarterly earnings per share were calculated for each company in the sample and straight lines were fitted to the quarterly earnings per share for each firm for consecutive eight quarter periods. Those firms whose earnings trend for a given eight-quarter period had a correlation coefficient in excess of .70 were selected for further analysis. For these firms an estimate of earnings was made for the ninth quarter. Then, if the firm's actual earnings were more than 1.5 standard errors of the estimate in excess of the estimated earnings, that firm's common stock price performance was measured by calculating a six-month price relative for the period beginning with the end of the second month of the tenth quarter.

A total of 211 common stocks were selected for the measurement of price performance. Of these, 136 had six-month price relatives greater than the corresponding price relatives for Standard and Poor's Industrial Index. By using a non-parametric sign test, Jones and Litzenberger concluded that this number of stocks outperforming the given market index was statistically significant at the .01 level. Furthermore, the average of the price relatives of the stock selected in each period exceeded the market index price relative in all ten of the periods examined. A sign test also showed that the selection technique used was significantly better than random selection at the .01 level. Based on these results Jones and Litzenberger concluded that there exists evidence that the stock market is not as perfect as some proponents of the random-walk theory had claimed in the sense that the market may not adjust instantaneously and correctly for every item of information that becomes available, thus giving rise to intermediate trends in stock prices.

Another study having implications for the random-walk theory was conducted by Ying (Ref. 21), to study the relationship between volumes of sales and stock prices. The data chosen for this investigation consisted of the daily closing price of Standard and Poor's 500 and the daily volume of stock sales on the New York Stock Exchange, both considered over the period from January 1957 to December 1962. The observations actually used were the price of the index

divided by the dividend rate and the volume of sales expressed as the ratio of the number of shares traded to the total number of shares outstanding. Using this data base the following variables were defined:

P(t) = closing price on the tth trading day

V(t) = volume of sales on the tth trading day

log P(t) = ln P(t)

log V(t) = ln V(t)

D log P(t) = log P(t) - log P(t-1)

D log V(t) = log V(t) - log V(t-1)

The first investigation conducted by Ying was to determine if there were interactions between the change in price and volume on the same day. To accomplish this a five-by-five contingency table was set up with cells corresponding to successive .2 quantiles of a normally distributed random variable. The rows corresponded to either log V(t) or D log V(t) and the columns corresponded to D log P(t). An independence-of-rows test was performed on both tables and the null hypothesis that the rows were independent was rejected at the .05 level of significance for both tables. Based on these results Ying reached the following conclusions:

1. A small volume was usually accompanied by either a fall or a small change in price.
2. A large volume was usually accompanied by a rise in price.
3. A large increase in volume was usually accompanied by either a large rise in price or a large fall in price.

The next investigation conducted was to study the effect of one or a combination of the variables D log P(t), D log V (t), and log V(t) on the distribution of D log P(t+1). To do this, the values of each of the three variables were placed into three classes which corresponded to the terciles of a normally distributed random variable. These classes were then used to form an estimated 27 by 27 probability transition matrix from time t to time t+1. Each row in the transition matrix was then used to form a cell in a three-way classification upon which an analysis of variance was performed. The analysis of variance yielded the following results:

1. The effect of D log P(t) on D log P(t+1) was significant at the .01 level.
2. The effect on D log V(t) on D log P(t+1) was not significant at the .05 level.
3. The effect of log V(t) on D log P(t+1) was significant at the .01 level.
4. All of the interaction effects of two factors were not significant at the .05 level.
5. The triple interaction effect was significant at the .01 level.

Based on these results, Ying reached the following conclusions:

1. A fall in price was followed, on the average, by a further fall in price.

2. A rise in price was followed, on the average, by a further rise in price.
3. A small volume was followed, on the average, by a fall in price.
4. A large volume was followed, on the average, by a rise in price.

Ying was of the opinion that conclusions one and two above cast serious doubt on the independence assumption of the random-walk hypothesis.

VI. SUMMARY

A general stock valuation model has been examined in which the value of a share of stock was defined to be a function of future earnings, rate of return, and future stock price. Models have also been presented which have attempted to evaluate each of these variables. It is felt that these various models are a fairly representative sample of the techniques of operations research as they have been applied to the stock market. It seems evident that this field has not nearly been exhausted and that the opportunity exists for a great deal more research to be undertaken. By pointing out work that has been done in the past, it is hoped that this thesis will stimulate interest in further work in the stock market on the part of persons with an operations research background.

Although not intended as a comprehensive list, the following areas are suggested as possible areas for future research and model-building:

1. Earnings and dividend estimation -- the literature does not have an over-abundance of articles concerning estimation of future earnings or dividends, and given that this is a key variable in any stock valuation model, it would seem that more work in this area would be useful.

2. Risk and rates of return -- although the treatment of risk was not included in the scope of this thesis, it appears that a great deal more research is necessary to arrive at models which are successful in relating risk and rates of return, and including the concept of risk in stock valuation.
3. Price trend research -- due to the wide disagreement over the behavior of stock price series, this area is still open for further research. In particular, more research is indicated concerning the relationship between stock prices and other explanatory variables besides past price data.
4. Technical analyst procedures -- the technical or professional stock market analyst has often been given too little credit in scholarly writing on the stock market. Although many of the technical analyst's procedures may be based on intuition, hunches, and "stock market sense", it would seem that some interesting and perhaps profitable research could be devoted to building and testing models which attempt to give a mathematical foundation to intuition.
5. "Success" analysis -- another area that would perhaps be fruitful would be an analysis of successful stocks in an attempt to relate the success of these stocks to a set of explanatory variables.
6. Portfolio analysis -- although this topic was not included in the scope of this thesis, it is a subject

that has received a considerable amount of attention in the stock market literature, and together with the entire spectrum of institutional investment constitutes a field with a great potential for further research.

LIST OF REFERENCES

1. Robichek, A. A., "Risk and the Value of Securities", Journal of Financial and Quantitative Analysis, v. 4, P. 513-538, December 1969.

2. Myers, S. C., "A Time-State-Preference Model of Security Valuation", Journal of Financial and Quantitative Analysis, v. 3, p. 1-34, March 1968.

3. Sloane, W. R. and Reisman, A., "Stock Evaluation Theory: Classification, Reconciliation, and General Model", Journal of Financial and Quantitative Analysis, v. 3, p. 171-204, June 1968.

4. Elton, E. J. and Gruber, M. J., "Earnings Estimates and the Accuracy of Expectational Data", Management Science, v. 18, p. B409-B424, April 1972.

5. Brown, P. and Niederhoffer, V., "The Predictive Content of Quarterly Earnings", The Journal of Business, v. 41, p. 488-497, October 1968.

6. Soldofsky, R. M. and Biderman, R., "Yield-Risk Measurements of the Performance of Common Stocks", Journal of Financial and Quantitative Analysis, v. 3, p. 59-74, March 1968.

7. Nerlove, M., "Factors Affecting Differences Among Rates of Return on Investments in Individual Common Stocks", The Review of Economics and Statistics, v. 50, p. 312-331, August 1968.

8. Benishay, H., "Variability in Earnings-Price Ratios", American Economic Review, v. 51, p. 81-94, March 1961.

9. Mandelbrot, B. and Taylor, H. M., "On the Distribution of Stock Price Differences", Operations Research, v. 15, p. 1057-1062, September-October 1967.

10. Smidt, S., "A New Look at the Random Walk Hypothesis", Journal of Financial and Quantitative Analysis, v. 3, p. 235-261, September 1968.

11. Seelenfreund, A., Parker, G. G. C., and Van Horne, J. C., "Stock Price Behavior and Trading", Journal of Financial and Quantitative Analysis, v. 3, p. 263-281, September 1968.

12. Jones, C. P. and Litzenberger, R. H., "Quarterly Earnings Reports and Intermediate Stock Price Trends", The Journal of Finance, v. 25, p. 143-148, March 1970.

13. Granger, C. W. J., "Some Aspects of the Random Walk Model of Stock Market Prices", International Economic Review, v. 9, p. 253-259, June 1968.

14. Osborne, M. F. M., "Brownian Motion in the Stock Market", Operations Research, v. 7, p. 145-173, March-April 1959.

15. Press, S. J., "A Compound Events Model for Security Prices," The Journal of Business, v. 40, p. 317-335, July 1967.

16. Kemp, A. G. and Reid, G. C., "The Random Walk Hypothesis and the Recent Behaviour of Equity Prices in Britain", Economica, v. 38, p. 28-51, February 1971.

17. Fama, E. F. and Blume, M. E., "Filter Rules and Stock Market Trading", The Journal of Business, v. 39, p. 226-241, January 1966.

18. James, F. E. Jr., "Monthly Moving Averages -- An Effective Investment Tool?", Journal of Financial and Quantitative Analysis, v. 3, p. 315-326, September 1968.

19. Alexander, S. S., "Price Movements in Speculative Markets: Trends or Random Walks, No. 2", Industrial Management Review, v. 5, p. 25-46, Spring 1964.

20. Mayor, T. H., "Short Trading Activities and the Price of Equities: Some Simulation Results", Journal of Financial and Quantitative Analysis, v. 3, p. 283-298, September 1968.

21. Ying, C. C., "Stock Market Prices and Volumes of Sales", Econometrica, v. 34, p. 676-685, July 1966.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
2.	Asst. Professor E. A. Brill, Code 55 Zg Department of Operations Research and Administrative Sciences Monterey, California 93940	1
3.	LCDR W. W. Wallace, SC, USN Fleet Material Support Office Mechanicsburg, Pennsylvania 17055	1
4.	Naval Postgraduate School Department of Operations Research and Administrative Sciences Monterey, California 93940	1
5.	Chief of Naval Personnel Pers 11b Department of the Navy Washington, D. C. 20370	1

INTERNALLY DISTRIBUTED
REPORT

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MODEL BUILDING IN THE STOCK MARKET: SOME EXAMPLES FROM THE LITERATURE		5. TYPE OF REPORT & PERIOD COVERED September 1973 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William Warren Wallace		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE September 1973 13. NUMBER OF PAGES 68
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

INTERNALLY DISTRIBUTED

REPORT

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

stock market

mathematical modelling

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The operations research technique of mathematical model building and testing as it has been applied in the field of stock market research is examined using various examples from the literature. The vehicle used for this examination is a general model for stock valuation. This general model illustrates that the value of a share of stock depends on estimates of future values of earnings or dividends, rates

DD FORM 1 JAN 73 1473 (Page 1)

EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)

of return, and future stock prices. Therefore models are illustrated which relate to each of these variables.

This examination of past applications of model building in the stock market is intended to stimulate interest in this area of research, and toward this end some areas for future research are suggested.

Thesis
W22256
c.1
Wallace
Model building in the stock market: some examples from the literature.
146747

Thesis
W22256
c.1
Wallace
Model building in the stock market: some examples from the literature.
146747

thesW22256
Model building in the stock market :
3 2768 000 99443 8
DUDLEY KNOX LIBRARY